

## Subquestion 2: Social Security Funding

Kylie King

This document explores the relationship between the approval for government spending on Social Security and public trust in government. Specifically, it aims to answer the following question.

### Subquestion 2

With the Social Security fund expected to be depleted by 2033,<sup>1</sup> has the proportion of people who think we are spending too little on social security increased in recent years? Does this correlate at all with the public trust scores?

### Data Collection

In this section, we initialize R packages and pull data downloaded from the General Society Survey (GSS) and Pew Research.

```
suppressPackageStartupMessages(library('tidyverse'))
```

The following data was a survey from GSS<sup>2</sup> where respondents were asked “Are we spending too much, too little, or about the right amount on Social Security?” Respondents could answer “ABOUT RIGHT”, “TOO LITTLE”, “TOO MUCH”, or “Do not know/Cannot Choose.” Other responses include Inapplicable, No answer, and Skipped on Web.

```
options(readr.show_col_types = FALSE)
ssc<- read_csv('data/socialsecurity.csv')
head(ssc)
```

```
# A tibble: 6 × 3
  year    id_natsoc
  <dbl> <dbl> <chr>
1 1972      1 Inapplicable
```

<sup>1</sup>Social Security Administration. (2025). Social Security Board of Trustees: Projection for Combined Trust Funds one year sooner than last year. Retrieved from <https://blog.ssa.gov/social-security-board-of-trustees-projection-for-combined-trust-funds-one-year-sooner-than-last-year/>

<sup>2</sup>Davern, Michael; Bautista, Rene; Freese, Jeremy; Herd, Pamela; and Morgan, Stephen L.; General Social Survey 1972-2024. [Machine-readable data file]. Principal Investigator, Michael Davern; Co-Principal Investigators, Rene Bautista, Jeremy Freese, Pamela Herd, and Stephen L. Morgan. Sponsored by National Science Foundation. NORC ed. Chicago: NORC, 2025: NORC at the University of Chicago [producer and distributor]. Data accessed from the GSS Data Explorer website at gssdataexplorer.norc.org. ‘natsoc’ variable accessed at <https://gssdataexplorer.norc.org/variables/193/vshow?back=variableList>.

```
2 1972      2 Inapplicable
3 1972      3 Inapplicable
4 1972      4 Inapplicable
5 1972      5 Inapplicable
6 1972      6 Inapplicable
```

```
print(unique(ssc$nat soc))
```

```
[1] "Inapplicable"                  "ABOUT RIGHT"
[3] "TOO LITTLE"                   "Do not Know/Cannot Choose"
[5] "TOO MUCH"                     "No answer"
[7] "Skipped on Web"
```

The following data is a collection of public survey responses synthesized by Pew Research Center<sup>3</sup> to evaluate trust in government. It is made up of the percentages of people who say they ‘trust the government to do what is right just about always/most of the time’ in each survey.

```
pct_public_trust <- read_csv('data/public_trust_pew.csv')
head(pct_public_trust)
```

```
# A tibble: 6 × 4
  Date       .     `Individual polls` `Moving average`
  <chr>     <chr>        <dbl>          <dbl>
1 5/19/2024 PEW            22             22
2 6/11/2023 PEW            16             19
3 5/1/2022 PEW            20             20
4 4/11/2021 PEW            24             21
5 8/2/2020 PEW            20             24
6 4/12/2020 PEW            27             21
```

## Data Cleaning

In this section, we will clean the data that was pulled in the previous section, to prepare it for Exploratory Data Analysis (EDA) and hypothesis testing.

### SSC Cleaning

To streamline this data, we will remove the answers Inapplicable, Skipped on Web, and No answer. This will ensure we capture all relevant responses that could be applicable to our question.

---

<sup>3</sup>Pew Research Center, National Election Studies, Gallup, ABC/Washington Post, CBS/New York Times, and CNN surveys. Data from 2020 and later comes from Pew Research Center’s online American Trends Panel; prior data is from telephone surveys. Data collected from <https://www.pewresearch.org/politics/2024/06/24/public-trust-in-government-1958-2024/>.

```
ssc_clean <- ssc[!(ssc$natsoc %in% c("Inapplicable", "No answer", "Skipped on Web")), ]  
print(unique(ssc_clean$natsoc))
```

```
[1] "ABOUT RIGHT"           "T00 LITTLE"  
[3] "Do not Know/Cannot Choose" "T00 MUCH"
```

```
cat(paste("Responses in original data:", nrow(ssc), "\nResponses remaining in  
clean data:", nrow(ssc_clean)))
```

```
Responses in original data: 75699  
Responses remaining in clean data: 57764
```

After removing the irrelevant answers, we retained 76.3% (57,764 responses) of the original data, all of which is now usable for the hypothesis tests. To get a better view of the data, and prepare it for other functions, we will aggregate the counts of survey responses by year.

```
yearly_counts <- ssc_clean |> group_by(year, natsoc) |> summarise(count =  
n(), .groups = 'drop')  
head(yearly_counts)
```

```
# A tibble: 6 × 3  
  year   natsoc       count  
  <dbl> <chr>      <int>  
1 1984 ABOUT RIGHT     344  
2 1984 Do not Know/Cannot Choose    46  
3 1984 T00 LITTLE      493  
4 1984 T00 MUCH        85  
5 1985 ABOUT RIGHT     586  
6 1985 Do not Know/Cannot Choose    46
```

## Public Trust Cleaning

This data has multiple entries for some survey years. For the purposes of this subquestion, we will take the average of each year's surveys, so that we have a single entry for each year. While this does cause some data loss, it will streamline the data for hypothesis testing. The data loss is minimized by taking the average for each year, instead of simply dropping the duplicate year values. We will also convert the percentages to proportions, which will allow better correlation between this data and the Social Security response data.

```
#ensuring uniform date format  
prop_public_trust <- pct_public_trust |> mutate(Date = mdy(Date))  
prop_public_trust <- prop_public_trust |> mutate(year = year(Date))
```

```
#aggregating by year to prevent double entries
prop_public_trust <- prop_public_trust |> group_by(year) |>
summarise(trust_prop = mean(`Moving average`, na.rm = TRUE) / 100, .groups =
"drop")
head(prop_public_trust)
```

```
# A tibble: 6 × 2
  year trust_prop
  <dbl>      <dbl>
1 1958       0.73
2 1964       0.77
3 1966       0.65
4 1968       0.62
5 1970       0.54
6 1972       0.53
```

## Exploratory Data Analysis

To prepare for our hypothesis tests, we will perform some EDA for each of our data. It is important to see what type of data we have, how it is distributed, and what patterns we might expect to see in the later hypothesis testing. We first initialize a theme so the charts are displayed uniformly.

```
eda_5100 <- theme_minimal() +
  theme(
    text = element_text(color = "black"),
    plot.title = element_text(size = 18, face = "bold", hjust = 0.5),
    axis.title = element_text(size = 14, face = "bold"),
    axis.text = element_text(size = 12),
    legend.title = element_text(size = 14, face = "bold"),
    legend.text = element_text(size = 12),
    strip.text.x = element_text(size = 14, face = "bold")
  )
theme_set(eda_5100)
```

We will first examine the Social Security survey responses for each year, to see if we can note any patterns.

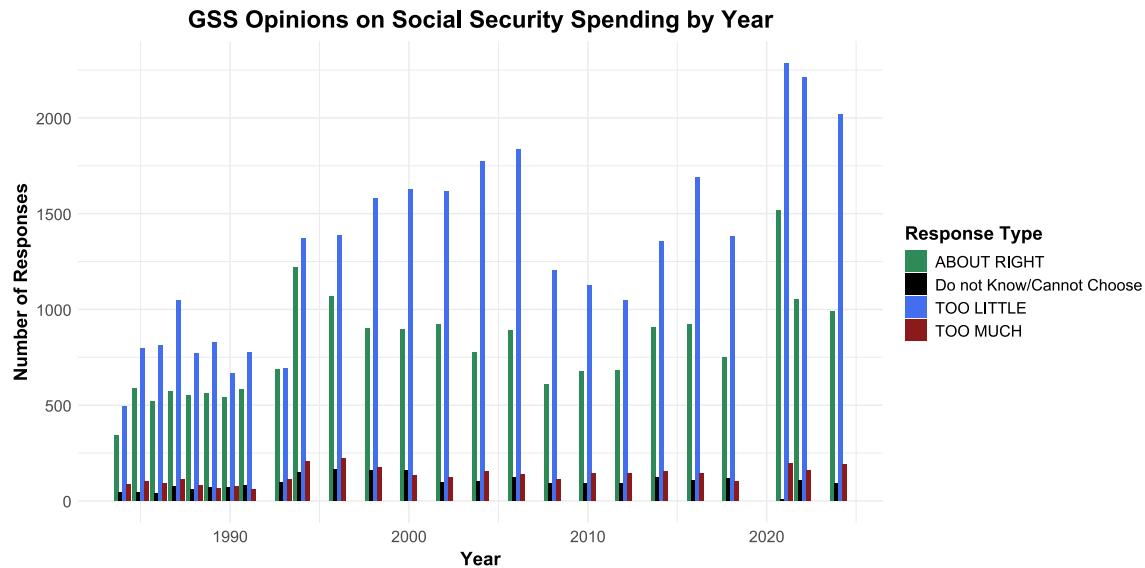
```
ssc_response_colors <- c('seagreen4', 'black','royalblue2','firebrick4')

ggplot(early_counts, aes(x = year, y = count, fill = natsoc)) +
  geom_col(position = "dodge") +
  labs(
    title = "GSS Opinions on Social Security Spending by Year",
    x = "Year",
    y = "Number of Responses",
```

```

fill = "Response Type"
) + scale_fill_manual(values = ssc_response_colors)

```



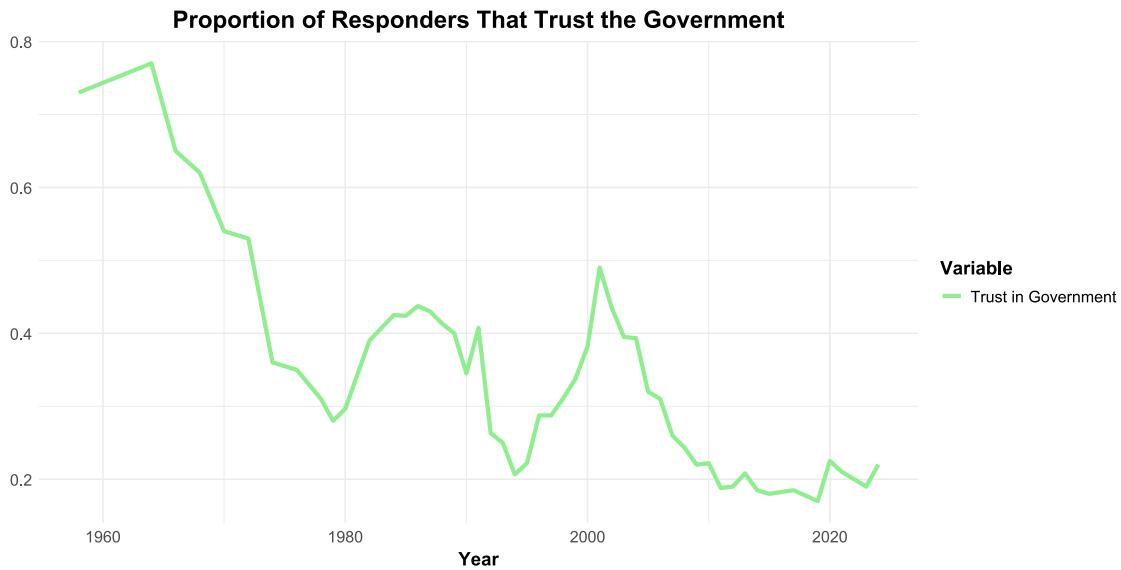
Note that the amount of people who think the government spends TOO LITTLE on Social Security is almost always higher than those who think the government spends ABOUT RIGHT, and far higher than those who think TOO MUCH or Cannot Choose. As a result, the rest of this analysis will focus on the proportion of people who think the government is spending TOO LITTLE on Social Security, and how that proportion changes each decade.

Next we will examine the Public Trust survey responses for each year.

```

ggplot(prop_public_trust, aes(x = year)) +
  geom_line(aes(y = trust_prop, color = "Trust in Government"), linewidth =
  1.5) +
  scale_color_manual(values = c("Trust in Government" = "lightgreen")) +
  labs(
    x = "Year",
    y = "",
    color = "Variable",
    title = "Proportion of Responders That Trust the Government"
  )

```



Note that the proportion of people who respond that they ‘trust the government to do what is right just about always/most of the time’ has had a general decline since the early 2000s, but is beginning to increase at the end of the study window, in 2024. While this data was collected over time, we will not be completing a time series test for this subquestion. The public trust value for each year will be compared to the proportion of people who think the government spends TOO LITTLE on Social Security for that year, but there will be no relation to past or future years.

Finally, we will display the potential relationship between the proportion of people who think the government spends TOO LITTLE on Social Security, and the proportion of people who ‘trust the government to do what is right just about always/most of the time’.

```
# isolating the proportions of people who think the gov spends TOO LITTLE on
# SSC
prop_tl_yearly <- ssc_clean |>
  mutate(too_little = ifelse(natsoc == "TOO LITTLE", 1, 0)) |>
  group_by(year) |>
  summarise(prop_too_little = mean(too_little), .groups = "drop")

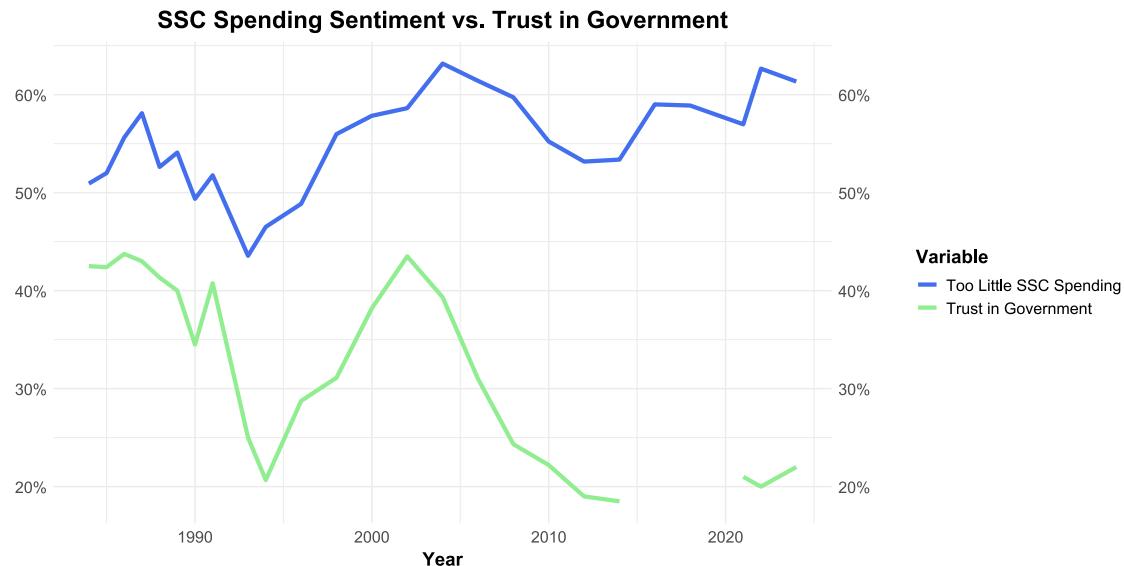
#joining the data
combined <- prop_tl_yearly |>
  left_join(prop_public_trust, by = "year")

ggplot(combined, aes(x = year)) +
  geom_line(aes(y = prop_too_little, color = "Too Little SSC Spending"),
            linewidth = 1.5) +
  geom_line(aes(y = trust_prop, color = "Trust in Government"), linewidth =
  1.5) +
  scale_y_continuous(labels = scales::percent, sec.axis = dup_axis()) +
  scale_color_manual(values = c("Too Little SSC Spending" = "royalblue2",
    "Trust in Government" = "green"))
```

```

"Trust in Government" = "lightgreen")) +
  labs(
    x = "Year",
    y = "",
    color = "Variable",
    title = "SSC Spending Sentiment vs. Trust in Government"
  )

```



Note that in the displayed time period of 1984-2010, the two variables appear to be somewhat correlated. However, after 2010 the proportion of people who think there is TOO LITTLE Social Security spending steadily increases, while the proportion of people who trust the government stays low. Also note that there are some missing values in the trust in government variable, as there are years where no survey was conducted.

## Hypothesis Testing

We will complete a Chi Squared Test for Independence to see whether the proportion of GSS respondents who think the government spends TOO LITTLE on Social Security is independent of the decade in which the response was submitted. This will help us identify whether a higher proportion of people in the past decade think that the government is spending TOO LITTLE on Social Security. After the Chi Squared Test, we will run a Pearson Correlation test to see if the proportion of GSS respondents who think the government spends TOO LITTLE on Social Security is at all correlated with the proportion of Pew Research respondents who ‘trust the government to do what is right just about always/most of the time’.

# Chi Squared Test of Independence

## Data Preparation

First we must aggregate our GSS survey responses into decade categories, and isolate the survey responses that think the government is spending TOO LITTLE on social security.

```
ssc_decade <- ssc_clean |>
  mutate(decade = case_when(
    year >= 1980 & year < 1990 ~ "1980s",
    year >= 1990 & year < 2000 ~ "1990s",
    year >= 2000 & year < 2010 ~ "2000s",
    year >= 2010 & year < 2020 ~ "2010s",
    year >= 2020 ~ "2020s"),
    too_little = ifelse(natsoc == "TOO LITTLE", 1, 0),
  )
```

## Assumptions

Next we must check our assumptions, to ensure we can run the Chi Squared Test.

- **Independence:** the GSS completes its surveys on new random representative samples of US citizens each year, so there is a low likelihood that a respondent was recorded more than once.<sup>4</sup>
- **Enough data per cell:** Because of the number of respondents, there should be >5 responses per cell in the table
- **Categorical variables:** The variables in this case are Decade and ‘Too Little’ or ‘Not Too Little’
- **Observations are counts:** This test uses the raw counts from the survey data, not the extracted proportions

This data fulfills all the assumptions, so we can move on to the Chi Squared Test for Independence.

## Hypotheses

**Null Hypothesis:** The proportion of respondents saying “TOO LITTLE” is the same in all decades. In other words, “decade” and “TOO LITTLE” response are independent.

**Alternative Hypothesis:** The proportion of respondents saying “TOO LITTLE” differs in at least one decade. That is, decade and response are not independent.

Alpha: 0.05

Now we will create a contingency table and run the Chi Squared Test of Independence.

```
# creating the contingency table
contingency_table <- table(ssc_decade$decade, ssc_decade$too_little)
contingency_table
```

---

<sup>4</sup>Frequently Asked Questions. (n.d.). Retrieved from <https://gss.norc.org/faq.html#accordion-2b38d5f6a8-9> General: How many people are interviewed for each GSS?

```
    0    1  
1980s 4004 4741  
1990s 6574 6471  
2000s 5326 8057  
2010s 5156 6598  
2020s 4318 6519
```

```
#running the chi squared test  
ssc_chi <- chisq.test(contingency_table)  
ssc_chi
```

Pearson's Chi-squared test

```
data: contingency_table  
X-squared = 399.82, df = 4, p-value < 2.2e-16
```

**Result from the raw test:** with a p value of 2.2e-16, which is less than our alpha of 0.05, we reject the null hypothesis that the two variables are independent. With a very large X-squared, we can also infer that there is a significant difference in the number of TOO LITTLE responses for at least one of the decades. However, we need to run a post-hoc residuals test to see which decade it is.

```
resid <- ssc_chi$stdres  
colnames(resid) <- c("Not Too Little", "Too Little")  
resid
```

	Not Too Little	Too Little
1980s	3.7885242	-3.7885242
1990s	16.8983437	-16.8983437
2000s	-11.0017369	11.0017369
2010s	-0.1664879	0.1664879
2020s	-9.5155700	9.5155700

## Results and Interpretation

The Chi Square test compares the real counts in each cell to the ‘expected’ counts in each cell if the variables were truly independent. This table is a comparison of how those real counts matched up to the ‘expected’ counts, or “did more people than would be expected think we are spending too little on Social Security?” It is also standardized, so the numbers do not directly correlate to counts.

If there is a negative value in the ‘Too Little’ column, that means that, for that decade, fewer people thought that there was too little spending on Social Security. If there is a positive value in the ‘Too Little’ column, many people thought that there was too little spending on Social Security. If the value is close to 0, there was not a significant sentiment about how much was being spent on Social Security. Any value +/- 2 is considered significant.

So with that in mind, here is the breakdown of the decades:

1980s: Significantly fewer people than expected thought there was Too Little spending  
1990s: Significantly fewer people than expected thought there was Too Little spending  
2000s: Significantly greater people than expected thought there was Too Little spending  
2010s: About the expected amount of people thought there was Too Little spending  
2020s: Significantly greater people than expected thought there was Too Little spending

Now that we know where to look, let's see if there is any correlation (or anticorrelation) between the proportion of people think the government is spending TOO LITTLE on Social Security, and the proportion of people who say they ‘trust the government to do what is right just about always/most of the time.’ We will run a Pearson Correlation test on the associated proportions.

```
# Pearson correlation (linear relationship)
cor.test(combined$prop_too_little, combined$trust_prop, method = "pearson")
```

```
Pearson's product-moment correlation

data: combined$prop_too_little and combined$trust_prop
t = 0.14477, df = 21, p-value = 0.8863
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
-0.3856454  0.4380765
sample estimates:
cor
0.03157629
```

With these results, it appears that the belief that Social Security spending is TOO LITTLE may not be correlated with the overall public trust in the government. However, because of our chi square test earlier, we do have statistically significant findings that there is a current, higher-than-average sentiment that the government is not spending enough on Social Security. This is a critical finding considering the recent news that the Social Security fund will run out by 2033. While this has not historically correlated with trust in the government, it does indicate that a significant number of people are worried about the current state of the SSC institution.

I want to refine this a little bit more and answer the following questions from the rubric, but had to move on to my second question for sake of time. I also need to break down the correlation results better.

Does the data answer your original question? How? Does the data help you defend against any objections? How? Are there any limitation on your conclusions, any angles you haven't considered? Is there a new result you found that you didn't expect to observe?